International Journal of Electrical Engineering and Technology (IJEET)
Volume 11, Issue 10, December 2020, pp. 382-391, Article ID: IJEET_11_10_049
Available online at https://iaeme.com/Home/issue/IJEET?Volume=11&Issue=10
ISSN Print: 0976-6545 and ISSN Online: 0976-6553
DOI: 10.34218/IJEET.11.10.2020.049

ENHANCING THE HUMAN EMOTION RECOGNITION WITH FEATURE EXTRACTION TECHNIQUES

Dr. P. Sumathy
Assistant Professor, Department of Computer Science, Bharathidasan University, Tiruchirappalli, Tamil Nadu, India

Ahilan Chandrasekaran
Research Scholar, Department of Computer Science, Bharathidasan University, Tiruchirappalli, Tamil Nadu, India

ABSTRACT

Human emotions are states of mental health that resolve spontaneously rather than through conscious exertion, and are accompanied by physiological changes in the facial muscles that signify expressions. Nonverbal communication methods such as expressions, eye movements, and gestures are used in many applications of human computer interaction. Identifying emotions is not an easy task because there is no difference between the emotions of a face, and there is also a lot of complexity and variability. The machine learning algorithm uses some open features to model the face. Automatic emotion recognition based on facial expression is an interesting research field, which has presented and applied in several areas such as safety, health and in human machine interfaces. Researches in this field are interested in developing techniques to interpret, code facial expressions and extract these features in order to have a better prediction by computer. Machine learning, one of the top emerging sciences, has an extensive range of applications. In this paper, the optimization techniques-based feature extraction techniques are used to enhance the recognition of the human emotion using facial images. The optimization techniques like Ant Colony Optimization, Particle Swarm Optimization, Genetic Algorithm are used. Various metrics are used to evaluate the performance of the feature extraction techniques for emotion recognition.

Keywords: Human Emotion Recognition, Facial Expression, Segmentation, Feature Extraction, Noise Removal, Ant Colony Optimization, Support Vector Machine

Cite this Article: P. Sumathy and Ahilan Chandrasekaran, Enhancing the Human Emotion Recognition with Feature Extraction Techniques, *International Journal of Electrical Engineering and Technology*, 11(10), 2020, pp. 382-391.
https://iaeme.com/Home/issue/IJEET?Volume=11&Issue=10

1. INTRODUCTION

An envisaged aim of artificial intelligence is to make the interaction between human and next generation computing systems more natural. In order to achieve efficient and smooth interaction between human and computer systems, a series of aspects of human behavior should be taken into account. One of the most important aspects concerns the emotional behavior and the affective state of the human. Next generation human-centered computing systems should possess the capacity to perceive, accurately analyze and deeply understand emotions as communicated by social and affective channels [1].

Emotions constitute an innate and important aspect of human behavior that colors the way of communication. Humans express their innate conditions through various channels, such as body language and facial expressions. Facial expressions are the most direct and meaningful channel of non-verbal communication, which forms a universal language of emotions that can instantly express a wide range of human emotional states, feelings and attitudes and assists in various cognitive tasks. The accurate analysis and interpretation of the emotional content of human facial expressions is essential for the deeper understanding of human behavior. Indeed, facial expressions are to wit the most cogent, naturally preeminent means for human beings to communicate emotions, comprehension, and intentions and to regulate interactions and communication with other people [1] [2].

Facial expressions considerably assist in direct communication and it has been indicated that during face-to-face human communication, 7% of the information is communicated by the linguistic part, such as the spoken words, 38% is communicated by paralanguage, such as the vocal part, and 55% is communicated by the facial expressions [3]. Indeed, even a simple signal such as a head nod or a smile can convey a large number of meanings [4] [5]. In general, facial expressions are the most natural, meaningful and important communication channel of human interaction and communication.

The recognition of facial expressions is assistive in a wide spectrum of systems and applications and is quite necessary for achieving naturalistic interaction. The facial expressions assist in various cognitive tasks; so, reading and interpreting the emotional content of human expressions is essential for deeper understanding of human condition. Therefore, the main aim of facial expression recognition methods and approaches is to enable machines to automatically estimate the emotional content of a human face. Giving computer applications the ability to recognize the emotional state of humans from their facial expressions is a very important and challenging task with wide ranging applications [6].

2. RELATED WORKS

Ngo, Quan T., and Seokhoon Yoon [6] applied deep learning techniques, and proposed a novel loss function called weighted-cluster loss, which is used during the fine-tuning phase. Specifically, the weighted-cluster loss function simultaneously improves the intra-class compactness and the inter-class separability by learning a class center for each emotion class. It also takes the imbalance in a facial expression dataset into account by giving each emotion class a weight based on its proportion of the total number of images.

Abdulrazaq, Maiwan B., et al [7] endeavors to inspect accuracy ratio of six classifiers based on Relief-F feature selection method, relying on the utilization of the minimum quantity of attributes. The classifiers in which the paper attempts to inspect are Multi-Layer Perceptron, Random Forest, Decision Tree, Support Vector Machine, K-Nearest Neighbor, and Radial Basis Function.

Liu, Xiaoqian, and Fengyu Zhou [8] applied a strategy of curriculum learning to facial expression recognition during the stage of training and a novel curriculum design method is

proposed. The system first employs the unsupervised density–distance clustering method to determine the clustering center of each category. Then, the dataset is divided into three subsets of various complexity according to the distance from each sample to the clustering center in the feature space. Importantly, the authors developed a multistage training process where a main model is trained by continuously adding harder samples to training set to increase the complexity.

Zheng, Hao, et al [9] proposed a discriminative DMTL (DDMTL) facial expression recognition method, which overcomes the above shortcomings by considering both the class label information and the samples' local spatial distribution information simultaneously. The authors further designed a siamese network to evaluate the local spatial distribution through an adaptive reweighting module, utilizing the class label information with different confidences.

Caroppo, Andrea, Alessandro Leone, and Pietro Siciliano [10] exploring the performance of existing deep architectures for the task of classifying expression of ageing adults are absent in the literature. In the present work a tentative to try this gap is done considering the performance of three recent deep convolutional neural networks models (VGG-16, AlexNet and GoogLeNet/Inception V1) and evaluating it on four different benchmark datasets (FACES, Lifespan, CIFE, and FER2013) which also contain facial expressions performed by elderly subjects. As the baseline, and with the aim of making a comparison, two traditional machine learning approaches based on handcrafted features extraction process are evaluated on the same datasets. Carrying out an exhaustive and rigorous experimentation focused on the concept of "transfer learning", which consists of replacing the output level of the deep architectures considered with new output levels appropriate to the number of classes (facial expressions), and training three different classifiers (i.e., Random Forest, Support Vector Machine and Linear Regression), VGG-16 deep architecture in combination with Random Forest classifier was found to be the best in terms of accuracy for each dataset and for each considered age-group.

Zhou, Linyi, et al [11] develop a 3D attention mechanism for feature refinement which selectively focuses on attentive channel entries and salient spatial regions of a convolution neural network feature map. Moreover, a deep metric loss termed Triplet-Center (TC) loss is incorporated to further enhance the discriminative power of the deeply-learned features with an expression-similarity constraint. It simultaneously minimizes intra-class distance and maximizes inter-class distance to learn both compact and separate features.

Liu, Yuanyuan, et al [12] proposed a dynamic multi-channel metric learning network for pose-aware and identity-invariant FER, called DML-Net, which can reduce the effects of pose and identity for robust FER performance. Specifically, DML-Net uses three parallel multi-channel convolutional networks to learn fused global and local features from different facial regions. Then it uses joint embedded feature learning to explore identity-invariant and pose-aware expression representations from fused region-based features in an embedding space. DML-Net is end-to-end trainable by minimizing deep multiple metric losses, FER loss, and pose estimation loss with dynamically learned loss weights, thereby suppressing overfitting and significantly improving recognition.

Najar, Fatma, et al [13] addressed the problem of human activities and facial expression recognition by investigating the effectiveness of Bayesian inference methods. Indeed, a novel method termed as Bayesian learning for finite multivariate generalized Gaussian mixture model is developed. The multivariate generalized Gaussian distribution is encouraged by its ability to model a large range of data and its shape flexibility. The authors developed a Markov Chain Monte Carlo within Metropolis-Hastings algorithm for proposed generative model. In this research, the authors tackled also some key issues related to machine learning and pattern recognition such as the statistical model's parameters estimation.

Xie, Yuan, et al [14] proposed a novel Adversarial Graph Representation Adaptation (AGRA) framework that unifies graph representation propagation with adversarial learning for cross-domain holistic-local feature co-adaptation. To achieve this, the authors first build a graph to correlate holistic and local regions within each domain and another graph to correlate these regions across different domains. Then, the authors learned the per-class statistical distribution of each domain and extract holistic-local features from the input image to initialize the corresponding graph nodes. Finally, the authors introduced two stacked graph convolution networks to propagate holistic-local feature within each domain to explore their interaction and across different domains for holistic-local feature co-adaptation.

Jiaming, Tang, et al [15] proposed a multi-scale and multi-region vector triangle texture feature extraction scheme based on weakly supervised clustering algorithm. According to the information gain rate of extracted features, combined with threshold selection and random dropout strategy, the best selection of vector triangle texture feature scale is explored, and the feature space is optimized under the premise of sufficient feature space information, the reduction of feature space is realized and the information redundancy is reduced. For the positive and negative expression units, the facial expression images in the data set are divided into two categories.

3. ANT COLONY OPTIMIZATION

The underlying metaphor of ant colony optimization (ACO) [16] is the way that some insects living in collaborative colonies look for food. Indeed, if an ant nest feels a food source, then some expeditions of ants go —by different paths— to search for this food, leaving a pheromone trail, a chemical substance that animals usually have, but very important for insects. This pheromone trail is an affective signal for other ants, that will recognize the way followed by its predecessors. Between all expeditions of ants, there will be some that arrive first to the food source because they took the shortest path, and then they will go back to the nest first than the other expeditions. Then, the shortest path has been reinforced in its pheromone trail; therefore, new expeditions will probably take that path more than others will, unless new better paths (or parts of paths) are found by some expeditions. It is expected that the pheromone trail of the shortest path is more and more intense, and the one of the other paths will evaporate.

When applying this principle to combinatorial optimization problems, we look for an implementation that uses the principle of reinforcement of good solutions, or parts of solutions, by the intensification of a value of "pheromone" that controls the probability of taking this solution or its part of solution. Now, this probability will depend not only on the pheromone value, but also on a value of a "local heuristic" or "short term vision", that suggests a solution or part of solution by a local optimization criterion, for example as greedy algorithms do.

An optimization method that uses ACO has at least the following components:

A representation, that enables the construction or modification of solutions by means of a probabilistic transition rule, that depends on the pheromone trail and the local heuristic.

- A local heuristic or visibility, noted η.
- An update rule for the pheromone, noted τ
- A probabilistic transition rule, that depends on η and τ.

4. PARTICLE SWARM OPTIMIZATION

It is a population-based stochastic optimization technique inspired by the social behavior of bird flocking. PSO was proposed by Eberhart and Kennedy in 1995 [17]. It is a metaheuristic as it can explore over a search space making no or few previous assumptions about the given problem and converges to an optimal solution. The candidate solutions, referred to as particles

in the technique, fly around in a multi-dimensional search space, to find out an optimal or sub-optimal solution by competition as well as by cooperation among them. Like GA, PSO is also initialized with a group of random particles and then it looks for optima through the movement of candidate solutions in the search space. Each particle is represented by a vector $x_i = (x_{i1}, x_{i2}, \dots, x_{iD})$ where D represents the number of features in the dataset. Each particle hence has a D-dimensional velocity represented as $v_i = (v_{i1}, v_{i2}, \dots, v_{iD})$. In every iteration, each particle is updated with three values: (1) previous velocity, which gives the trend of flow of the particles over the search space; (2) pbest, which gives the particles' best fitness values till the present iteration and (3) gbest, which gives the whole generation's best fitness value till the present iteration. The position and velocity of the particles are updated using the following equations:

$$v_{id}^{k+1} = w * v_{id}^{k} + c1 * r1 * \left(p_{id} - x_{id}^{k}\right) + c2 * r2 * \left(p_{gd} - x_{id}^{t}\right) \quad (1)$$

$$x_{id}^{k+1} = x_{id}^{k} + v_{id}^{k+1} \quad (2)$$

Here k represents the kth iteration and d represents the dth feature in the vector. w represents the inertia factor which assigns a weight to the impact of previous velocity. $c1$ and $c2$ are acceleration constants. $r1$ and $r2$ are random numbers in the range [0, 1]. p_{gd} and g_{id} denote the state of dth feature in pbest and gbest.

5. GENETIC ALGORITHM

GA is a popular evolutionary algorithm computational method developed by Holland in early 1975 and later enhanced by Goldberg [18]. It is a global search technique that solves a given problem by mimicking the natural process of evolution. Based on Darwin's theory, GA utilizes the concept of reproduction and survival of the fittest. GA exploits new and better solutions without any presumption such as continuity or unimodality. As a process, GA has large potential, and due to this, over the years GA has been used for designing, optimizing telecommunication, traffic and shipment routing, gaming, market and financial analysis and many more. The increase in its use in different sectors is because of the fact that GA can handle a large number of parameters, and it comes with a solution which is satisfying enough though may not be the best.

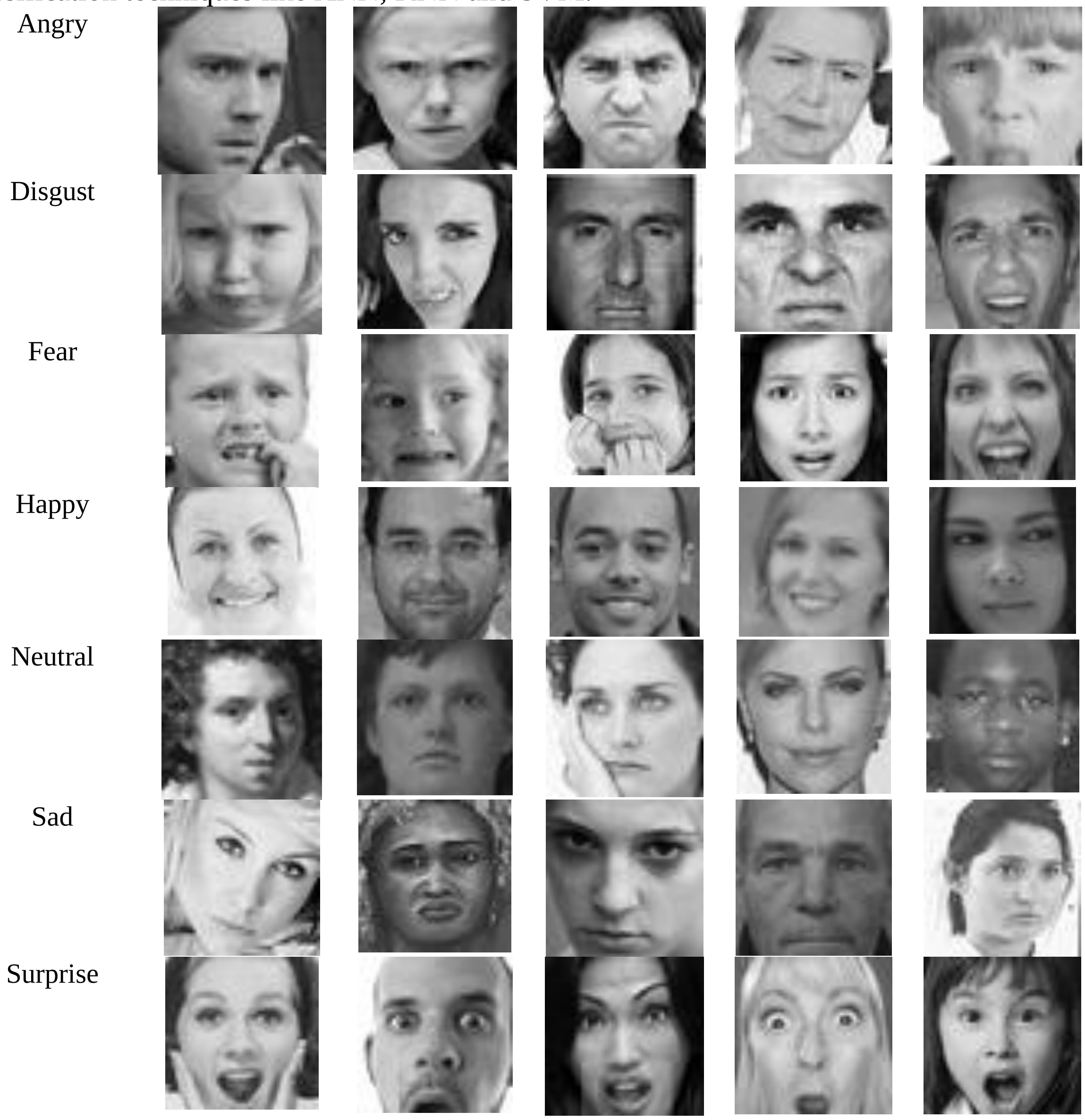
GA consists of a set of solutions, chromosomes or individuals which are strings of binary values, "0"s and "1"s. Each value ("0" or "1") determines the state of attributes in the chromosome. A set of such chromosomes is referred to as a population. Each chromosome is then evaluated using a fitness function. After ranking the chromosomes according to their fitness values, they undergo genetic operations such as crossover and mutation. For this, two chromosomes are selected on the basis of their positions on a roulette wheel (biased according to each chromosome's fitness). The two chromosomes first go through crossover and then mutation is applied to increase the local coverage of search space by the chromosomes, thereby decreasing the chances of being stuck at a local optimum. If the evolution process generates stronger offspring chromosomes than the previous ones, the algorithm replaces them. The evolution process repeats until it meets the end criteria.

6. RESULT AND DISCUSSION

6.1. Image Dataset

The face emotion recognition dataset is taken from the Kaggle repository [19]. The dataset is composed of angry, disgust, fear, happy, neutral, sad and surprise emotions. For this paper, 100 images from each emotion category are considered to evaluate the performance of the

Optimization based Feature Extraction techniques like PSO, GA and ACO using three classification techniques like ANN, KNN and SVM.



6.2. Performance Metrics

Table 1 depicts the performance metrics used in this research paper.

Table 1 Performance Metrics used in this paper

Performance Metrics	Equation
Detection Rate	$\frac{TP+TN}{TP+TN+FP+FN}$
Sensitivity	$\frac{TP}{TP+FN}$
Specificity	$\frac{TN}{TN+FP}$
False Positive Rate	1- Specificity
Miss Rate	1- Sensitivity

Table 2 depicts the Detection Rate (in %) obtained by the Optimization techniques-based feature extraction techniques using the classifiers like Artificial Neural Network (ANN), Convolutional Neural Network (CNN) and K-Nearest Neighbor (KNN). Figure 2 gives the graphical representation of the Detection Rate (in %) obtained by the Optimization techniques-based feature extraction techniques using the classifiers like ANN, CNN and KNN. From the table 2 and figure 2, it is clear that the ACO with CNN gives increased detection rate when it is compared with other feature extraction techniques.

Table 2 Detection Rate (in %) obtained by the Optimization based Feature Extraction techniques using CNN, ANN and KNN classifier

Feature Extraction Techniques	**Detection Rate obtained (in %) Classification Techniques**		
	CNN	ANN	KNN
PSO	55.67	53.62	47.41
GA	56.64	53.69	51.23
ACO	64.58	63.37	49.25

Figure 2: Graphical representation of the Detection Rate (in %) obtained by the Optimization based Feature Extraction techniques using CNN, ANN and KNN classifier

Table 3 depicts the Sensitivity (in %) obtained by the obtained by the Optimization techniques-based feature extraction techniques using the classifiers like CNN, ANN and KNN. Figure 3 depicts the graphical representation of the Sensitivity (in %) obtained by the obtained by the Optimization techniques-based feature extraction techniques using the classifiers like CNN, ANN and KNN. From the table 3 and figure 3, it is clear that the ACO with CNN gives increased sensitivity when it is compared with other feature extraction techniques.

Table 3 Sensitivity (in %) obtained by the obtained by the Optimization based Feature Extraction techniques using CNN, ANN and KNN classifier

Feature Extraction Techniques	**Sensitivity obtained by Classification Techniques**		
	CNN	ANN	KNN
PSO	53.34	51.68	49.26
GA	55.34	55.81	48.24
ACO	67.83	59.54	56.48

Figure 3: Graphical representation of the Sensitivity (in %) obtained by the Optimization based Feature Extraction techniques using CNN, ANN and KNN classifier

Table 4 depicts the Specificity (in %) obtained by the Optimization based Feature Extraction techniques using CNN, ANN and KNN classifier. Figure 4 depicts the graphical representation of the Specificity (in %) obtained by the Optimization based Feature Extraction techniques using CNN, ANN and KNN classifier. From the table 4 and figure 4, it is clear that the ACO with CNN gives increased specificity when it is compared with other feature extraction techniques.

Table 4 Specificity (in %) obtained by the Optimization based Feature Extraction techniques using CNN, ANN and KNN classifier

Feature Extraction Techniques	**Specificity obtained by Classification Techniques**		
	CNN	ANN	KNN
PSO	55.54	53.71	52.32
GA	56.45	51.68	47.81
ACO	63.76	58.58	55.82

Figure 4: Graphical representation of the Specificity (in %) obtained by the Optimization based Feature Extraction techniques using CNN, ANN and KNN classifier

Table 5 depicts the False Positive Rate (in %) obtained by the Optimization based Feature Extraction techniques using CNN, ANN and KNN classifier. Figure 5 depicts the graphical representation of the False Positive Rate (in %) obtained by the Optimization based Feature Extraction techniques using CNN, ANN and KNN classifier. From the table 5 and figure 5, it is clear that the ACO with CNN gives reduced FPR when it is compared with other feature extraction techniques.

Table 5 False Positive Rate (in %) obtained by the Optimization based Feature Extraction techniques using CNN, ANN and KNN classifier

Feature Extraction Techniques	**False Positive Rate obtained by Classification Techniques**		
	CNN	ANN	KNN
PSO	44.46	46.29	47.68
GA	43.55	48.32	52.19
ACO	36.24	41.42	44.18

Table 6 depicts the Miss Rate (in %) obtained by the Optimization based Feature Extraction techniques using CNN, ANN and KNN classifier. Figure 6 depicts the graphical representation of the Miss Rate (in %) obtained by Optimization based Feature Extraction techniques using CNN, ANN and KNN classifier. From the table 6 and figure 6, it is clear that the ACO with CNN gives reduced miss rate when it is compared with other feature extraction techniques.

Table 6 Miss Rate (in %) obtained by the Optimization based Feature Extraction techniques using CNN, ANN and KNN classifier

Feature Extraction Techniques	**Miss Rate obtained by Classification Techniques**		
	CNN	ANN	KNN
PSO	46.66	48.32	50.74
GA	44.66	44.19	51.76
ACO	32.17	40.46	43.52

Figure 6: Graphical representation of the Miss Rate (in %) obtained by the Optimization based Feature Extraction techniques using CNN, ANN and KNN classifier

7. CONCLUSION

The accurate analysis and interpretation of the emotional content of human facial expressions is essential for deeper understanding human behavior. Although a human can detect and interpret faces and facial expressions naturally, with little or no effort, accurate and robust facial expression recognition by computer systems is still a great challenge. Through an effective feature extraction technique, different facial expression can be easily classified into their appropriate class. In this research paper, the optimization based feature extraction techniques are utilized to enhance the classification of the human facial expression. From the result and discussion, the ACO increased the detection rate, specificity, and sensitivity with CNN classifier and also it reduced the false positive rate and miss rate with CNN than other classification techniques. The performance of the ACO based feature extraction is better when it is compared with the other optimization-based feature extraction techniques like PSO and GA.

REFERENCES

[1] Umer, Saiyed, et al. "Facial expression recognition with trade-offs between data augmentation and deep learning features." *Journal of Ambient Intelligence and Humanized Computing* (2021): 1-15.

[2] Fan, Yingruo, Victor Li, and Jacqueline CK Lam. "Facial expression recognition with deeply-supervised attention network." *IEEE transactions on affective computing* (2020).

[3] Wu, Shaohuai, and Bo Wang. "Facial expression recognition based on computer deep learning algorithm: taking cognitive acceptance of college students as an example." *Journal of Ambient Intelligence and Humanized Computing* (2021): 1-12.

[4] Abdullah, Sharmeen M. Saleem Abdullah, et al. "Multimodal emotion recognition using deep learning." *Journal of Applied Science and Technology Trends* 2.02 (2021): 52-58.

[5] Kim, Seong-Kyu Steve. "Affective Computing Among Individuals in Deep Learning." *Journal of Multimedia Information System* 7.2 (2020): 115-124.

[6] Ngo, Quan T., and Seokhoon Yoon. "Facial Expression Recognition Based on Weighted-Cluster Loss and Deep Transfer Learning Using a Highly Imbalanced Dataset." *Sensors* 20.9 (2020): 2639.

[7] Abdulrazaq, Maiwan B., et al. "An analytical appraisal for supervised classifiers' performance on facial expression recognition based on relief-F feature selection." *Journal of Physics: Conference Series*. Vol. 1804. No. 1. IOP Publishing, 2021.

[8] Liu, Xiaoqian, and Fengyu Zhou. "Improved curriculum learning using SSM for facial expression recognition." *The Visual Computer* 36.8 (2020): 1635-1649.

[9] Zheng, Hao, et al. "Discriminative deep multi-task learning for facial expression recognition." *Information Sciences* 533 (2020): 60-71.

[10] Caroppo, Andrea, Alessandro Leone, and Pietro Siciliano. "Comparison between deep learning models and traditional machine learning approaches for facial expression recognition in ageing adults." *Journal of Computer Science and Technology* 35.5 (2020): 1127-1146.

[11] Zhou, Linyi, et al. "Discriminative attention-augmented feature learning for facial expression recognition in the wild." *Neural Computing and Applications* (2021): 1-12.

[12] Liu, Yuanyuan, et al. "Dynamic multi-channel metric network for joint pose-aware and identity-invariant facial expression recognition." *Information Sciences* 578 (2021): 195-213.

[13] Najar, Fatma, et al. "Efficient Statistical Learning Framework with Applications to Human Activity and Facial Expression Recognition." *2020 5th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*. IEEE, 2020.

[14] Xie, Yuan, et al. "Adversarial graph representation adaptation for cross-domain facial expression recognition." *Proceedings of the 28th ACM international conference on Multimedia*. 2020.

[15] Jiaming, Tang, et al. "Texture feature extraction and optimization of facial expression based on weakly supervised clustering." *Systems Science & Control Engineering* 9.1 (2021): 514-528.

[16] Adeli, Ali, and Ali Broumandnia. "Image steganalysis using improved particle swarm optimization based feature selection." *Applied Intelligence* 48.6 (2018): 1609-1622.

[17] Zhang, Shike, et al. "Prediction of rock size distribution in mine bench blasting using a novel ant colony optimization-based boosted regression tree technique." *Natural Resources Research* 29.2 (2020): 867-886.

[18] Alawad, Duaa Mohammad, Avdesh Mishra, and Md Tamjidul Hoque. "AIBH: accurate identification of brain hemorrhage using genetic algorithm based feature selection and stacking." *Machine Learning and Knowledge Extraction* 2.2 (2020): 56-77.

[19] https://www.kaggle.com/jonathanoheix/face-expression-recognition-dataset/data#